With this lemma a Lyupunov function $V = x^{\mathrm{T}}Px$ can be constructed such that P satisfies not only the Lyapunov equation (D.15) but also the input-output condition (D.16) from which the restriction to relative degree zero ($d > 0$) or one ($d = 0$, $cb > 0$) is apparent. The main utility of this special Lyapunov function for adaptive and cascade designs is that the indefinite term in its derivative $\dot{V}$ depends on the output y and not on the whole state x.

Appendix E

Parameter Projection

The modular adaptive controllers have a point of singularity $\hat{b}_m = 0$, where $\hat{b}_m$ is the estimate of the high-frequency gain (virtual control coefficient) b_m. In order to prevent $\hat{b}_m$ from taking the value zero, we use the parameter projection in our identifiers. For this, we need to know the sign of the actual high-frequency gain b_m. We first give a treatment of projection for a general convex parameter set and then specialize to the case where only the high-frequency gain is constrained.

Let us define the following convex set

$$\Pi = \left\{ \hat{\theta} \in \mathbb{R}^p \,\middle|\, \mathcal{P}(\hat{\theta}) \leq 0 \right\}, \tag{E.1}$$

where by assuming that the convex function $\mathcal{P} : \mathbb{R}^p \to \mathbb{R}$ is smooth, we assure that the boundary $\partial\Pi$ of Π is smooth. Let us denote the interior of Π by $\mathring{\Pi}$ and observe that $\nabla_{\hat{\theta}}\mathcal{P}$ represents an outward normal vector at $\hat{\theta} \in \partial\Pi$. The standard projection operator is

$$\mathrm{Proj}\{\tau\} = \begin{cases} \tau, & \hat{\theta} \in \mathring{\Pi} \ \text{ or } \nabla_{\hat{\theta}}\mathcal{P}^{\mathrm{T}}\tau \leq 0 \\ \left(I - \Gamma \dfrac{\nabla_{\hat{\theta}}\mathcal{P}\ \nabla_{\hat{\theta}}\mathcal{P}^{\mathrm{T}}}{\nabla_{\hat{\theta}}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat{\theta}}\mathcal{P}} \right)\tau, & \hat{\theta} \in \partial\Pi \text{ and } \nabla_{\hat{\theta}}\mathcal{P}^{\mathrm{T}}\tau > 0\,, \end{cases} \tag{E.2}$$

where Γ belongs to the set $\mathcal{G}$ of all positive definite symmetric $p \times p$ matrices. Although Proj is a function of three arguments, τ, $\hat{\theta}$ and Γ, for compactness of notation we write only $\mathrm{Proj}\{\tau\}$.

The meaning of (E.2) is that, when $\hat{\theta}$ is in the interior of Π or at the boundary with τ pointing inward, then $\mathrm{Proj}\{\tau\} = \tau$. When $\hat{\theta}$ is at the boundary with τ pointing outward, then Proj projects τ on the hyperplane tangent to $\partial\Pi$ at $\hat{\theta}$.

In general, the mapping (E.2) is discontinuous. This is undesirable for two reasons. First, the discontinuity represents a difficulty for implementation in continuous time. Second, since the Lipschitz continuity is violated, we cannot use standard theorems for existence of solutions. Therefore, we need to smooth the projection operator. Let us consider the following convex set

$$\Pi_\varepsilon = \left\{ \hat\theta \in \mathbb{R}^p \,\middle|\, \mathcal{P}(\hat\theta) \le \varepsilon \right\}, \tag{E.3}$$

which is a union of the set Π and an $O(\varepsilon)$-boundary layer around it. We now modify (E.2) to achieve continuity of the transition from the vector field τ on the boundary of Π to the vector field $\left(I - \Gamma \frac{\nabla_{\hat\theta}\mathcal{P}\, \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}}{\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat\theta}\mathcal{P}}\right)\tau$ on the boundary of Π_ε:

3rd Algorithm

$$\mathrm{Proj}\{\tau\} = \begin{cases} \tau, & \hat\theta \in \mathring{\Pi} \text{ or } \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau \le 0 \\ \left(I - c(\hat\theta)\Gamma \dfrac{\nabla_{\hat\theta}\mathcal{P}\, \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}}{\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat\theta}\mathcal{P}}\right)\tau, & \hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi} \text{ and } \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau > 0 \end{cases} \tag{E.4}$$

$$c(\hat\theta) = \min\left\{1, \frac{\mathcal{P}(\hat\theta)}{\varepsilon}\right\}. \tag{E.5}$$

It is helpful to note that $c(\partial\Pi) = 0$ and $c(\partial\Pi_\varepsilon) = 1$.

In the proofs of stability of identifiers we need the following technical properties of the projection operator (E.4).

Lemma E.1 (Projection Operator) *The following are the properties of the projection operator (E.4):*

(i) The mapping $\mathrm{Proj}: \mathbb{R}^p \times \Pi_\varepsilon \times \mathcal{G} \to \mathbb{R}^p$ *is locally Lipschitz in its arguments* $\tau, \hat\theta, \Gamma$.

(ii) $\mathrm{Proj}\{\tau\}^{\mathrm{T}}\Gamma^{-1}\mathrm{Proj}\{\tau\} \le \tau^{\mathrm{T}}\Gamma^{-1}\tau, \quad \forall\hat\theta \in \Pi_\varepsilon$.

(iii) Let $\Gamma(t), \tau(t)$ *be continuously differentiable and*

$$\dot{\hat\theta} = \mathrm{Proj}\{\tau\}, \qquad \hat\theta(0) \in \Pi_\varepsilon.$$

Then, on its domain of definition, the solution $\hat\theta(t)$ *remains in* Π_ε.

(iv) $-\tilde\theta^{\mathrm{T}}\Gamma^{-1}\mathrm{Proj}\{\tau\} \le -\tilde\theta^{\mathrm{T}}\Gamma^{-1}\tau, \quad \forall\hat\theta \in \Pi_\varepsilon, \theta \in \Pi$.

Proof. *(i)* The proof of this point is lengthy but straightforward. The reader is referred to [157, Lemma (103)].

(ii) For $\hat\theta \in \mathring{\Pi}$ or $\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau \le 0$, we have $\mathrm{Proj}\{\tau\} = \tau$ and *(ii)* trivially holds with equality. Otherwise, a direct computation gives

$$\begin{aligned} \mathrm{Proj}\{\tau\}^{\mathrm{T}}\Gamma^{-1}\mathrm{Proj}\{\tau\} &= \tau^{\mathrm{T}}\Gamma^{-1}\tau - 2c(\hat\theta)\frac{\left(\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau\right)^2}{\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat\theta}\mathcal{P}} + c(\hat\theta)^2\frac{\left|\nabla_{\hat\theta}\mathcal{P}\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau\right|_\Gamma^2}{\left(\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat\theta}\mathcal{P}\right)^2} \\ &= \tau^{\mathrm{T}}\Gamma^{-1}\tau - c(\hat\theta)\left(2 - c(\hat\theta)\right)\frac{\left(\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau\right)^2}{\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat\theta}\mathcal{P}} \\ &\le \tau^{\mathrm{T}}\Gamma^{-1}\tau, \end{aligned} \tag{E.6}$$

where the last inequality follows by noting that $c(\hat\theta) \in [0,1]$ for $\hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi}$.

(iii) Using the definition of the Proj operator, we get

$$\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\mathrm{Proj}\{\tau\} = \begin{cases} \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau, & \hat\theta \in \mathring{\Pi} \text{ or } \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau \le 0 \\ \left(1 - c(\hat\theta)\right)\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau, & \hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi} \text{ and } \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau > 0, \end{cases} \tag{E.7}$$

which, in view of the fact that $c(\hat\theta) \in [0,1]$ for $\hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi}$, implies that

$$\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\mathrm{Proj}\{\tau\} \le 0 \text{ whenever } \hat\theta \in \partial\Pi_\varepsilon, \tag{E.8}$$

that is, the vector $\mathrm{Proj}\{\tau\}$ either points inside Π_ε or is tangential to the hyperplane of $\partial\Pi_\varepsilon$ at $\hat\theta$. Since $\hat\theta(0) \in \Pi_\varepsilon$, it follows that $\hat\theta(t) \in \Pi_\varepsilon$ as long as the solution exists.

(iv) For $\hat\theta \in \mathring{\Pi}$, *(iv)* trivially holds with equality. For $\hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi}$, since $\theta \in \Pi$ and $\mathcal{P}$ is a convex function, we have

$$(\theta - \hat\theta)^{\mathrm{T}}\nabla_{\hat\theta}\mathcal{P} \le 0 \text{ whenever } \hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi}. \tag{E.9}$$

With (E.9) we now calculate

$$\begin{aligned} -\tilde\theta^{\mathrm{T}}\Gamma^{-1}\mathrm{Proj}\{\tau\} &= -\tilde\theta^{\mathrm{T}}\Gamma^{-1}\tau \\ &\quad + \begin{cases} 0, & \hat\theta \in \mathring{\Pi} \text{ or } \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau \le 0 \\ c(\hat\theta)\dfrac{(\tilde\theta^{\mathrm{T}}\nabla_{\hat\theta}\mathcal{P})(\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau)}{\nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\Gamma\nabla_{\hat\theta}\mathcal{P}}, & \hat\theta \in \Pi_\varepsilon \setminus \mathring{\Pi} \text{ and } \nabla_{\hat\theta}\mathcal{P}^{\mathrm{T}}\tau > 0 \end{cases} \\ &\le -\tilde\theta^{\mathrm{T}}\Gamma^{-1}\tau, \end{aligned} \tag{E.10}$$

which completes the proof. □

Since we intend to use the projection operator only to keep the estimate $\hat b_m$ of the high-frequency gain b_m from becoming zero, we now specialize the projection operator for this case. We assume that $|b_m| \ge \varsigma_m > 0$, where $\operatorname{sgn} b_m$ and ς_m are known. Recalling that b_m is the first element of the parameter vector θ, i.e., $\theta = [b_m, \theta_2, \ldots, \theta_p]^{\mathrm{T}}$, we define $\mathcal{P}(\hat\theta) = \varsigma_m - \hat b_m \operatorname{sgn} b_m$ and note that $\nabla_{\hat\theta}\mathcal{P} = -\operatorname{sgn} b_m e_1^{\mathrm{T}}$. Let us denote the nominal vector field for the parameter update law by $\tau = [\tau_1, \tau_2, \ldots, \tau_p]^{\mathrm{T}}$ and choose $\varepsilon \in (0, \varsigma_m)$. The update law of the form $\dot{\hat\theta} = \mathrm{Proj}\{\tau\}$ using the projection operator (E.4)–(E.5) is given by

$$\dot{\hat b}_m = \tau_1 \begin{cases} 1, & \hat b_m \operatorname{sgn} b_m > \varsigma_m \text{ or } \tau_1 \operatorname{sgn} b_m \ge 0 \\ \max\left\{0, \frac{\varepsilon - \varsigma_m + \hat b_m \operatorname{sgn} b_m}{\varepsilon}\right\}, & \hat b_m \operatorname{sgn} b_m \le \varsigma_m \text{ and } \tau_1 \operatorname{sgn} b_m < 0 \end{cases} \tag{E.11}$$

$$\dot{\hat\theta}_i = \begin{cases} \tau_i, & \hat b_m \operatorname{sgn} b_m > \varsigma_m \text{ or } \tau_1 \operatorname{sgn} b_m \ge 0 \\ \tau_i - \tau_1 \frac{\Gamma_{1i}}{\Gamma_{11}} \min\left\{1, \frac{\varsigma_m - \hat b_m \operatorname{sgn} b_m}{\varepsilon}\right\}, & \hat b_m \operatorname{sgn} b_m \le \varsigma_m \text{ and } \tau_1 \operatorname{sgn} b_m < 0 \end{cases} \tag{E.12}$$

where $i = 2, \ldots, p$, and Γ_{1i} is the $(1, i)$ element of the positive definite symmetric matrix Γ. This update law achieves $\hat{b}_m(t) \operatorname{sgn} b_m \geq \varsigma_m - \varepsilon, \forall t \geq 0$ whenever $\hat{b}_m(0) \operatorname{sgn} b_m \geq \varsigma_m - \varepsilon$.

From expression (E.12) one should observe that when the update law is gradient it can be simplified to $\dot{\hat{\theta}}_i = \tau_i$ by the choice of the adaptation gain matrix $\Gamma = \operatorname{diag}\{\gamma_1, \Gamma_2\}$, where γ_1 is a positive scalar and Γ_2 is a symmetric positive definite $(p-1) \times (p-1)$ matrix. Thus, projection is applied only to $\dot{\hat{b}}_m$. This simplification is not possible with the least-squares update law because $\Gamma(t)$ is not a constant matrix, so it does not maintain a block-diagonal structure even if $\Gamma(0)$ is block-diagonal.

By setting $\varepsilon = 0$ in (E.11) and (E.12), we obtain update laws with the standard discontinuous projection, which in the case of gradient estimation simplifies to

$$\dot{\hat{b}}_m = \begin{cases} \tau_1 & \hat{b}_m \operatorname{sgn} b_m > \varsigma_m \text{ or } \tau_1 \operatorname{sgn} b_m \geq 0 \\ \\ 0 & \hat{b}_m \operatorname{sgn} b_m \leq \varsigma_m \text{ and } \tau_1 \operatorname{sgn} b_m < 0 \end{cases} \tag{E.13}$$

For the most part, the properties of the projection operator given in this appendix are recapitulated from [157] (see also [43, 51, 165]).

Appendix F

Nonlinear Swapping

The well-known Swapping Lemma [138] is ubiquitous in adaptive linear control. Here we provide its nonlinear counterpart.

Lemma F.1 (Nonlinear Swapping Lemma) *Consider the nonlinear time-varying system*

$$\Sigma_1 : \quad \begin{aligned} \dot{z} &= A(z,t)z + g(z,t)W(z,t)^{\mathrm{T}}\tilde{\theta} - Q(z,t)^{\mathrm{T}}\dot{\tilde{\theta}} + E(z,t)e \\ y_1 &= h(z,t)z + l(z,t)W(z,t)^{\mathrm{T}}\tilde{\theta} \end{aligned} \tag{F.1}$$

where $\tilde{\theta} : \mathbb{R}_+ \to \mathbb{R}^p$ is differentiable, $e : \mathbb{R}_+ \to \mathbb{R}$ is continuous and $\lim_{t\to\infty} e(t) = 0$, the matrix-valued functions $A : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{n\times n}$, $g : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{n\times m}$, $W : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{p\times m}$, $Q : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{p\times n}$, $E : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$, and $l : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{r\times m}$ are locally Lipschitz in z and continuous and bounded in t, and $h : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^{r\times n}$ is bounded in z and t. Along with (F.1) consider the linear time-varying systems

$$\Sigma_2 : \quad \begin{aligned} \dot{\Omega}^{\mathrm{T}} &= A(z,t)\Omega^{\mathrm{T}} + g(z,t)W(z,t)^{\mathrm{T}} \\ y_2 &= h(z,t)\Omega^{\mathrm{T}} + l(z,t)W(z,t)^{\mathrm{T}} \end{aligned} \tag{F.2}$$

$$\Sigma_3 : \quad \begin{aligned} \dot{\psi} &= A(z,t)\psi + \Omega^{\mathrm{T}}\dot{\tilde{\theta}} + Q(z,t)^{\mathrm{T}}\dot{\tilde{\theta}} \\ y_3 &= -h(z,t)\psi\,. \end{aligned} \tag{F.3}$$

Assume that $z(t)$ is continuous on $[0,\infty)$ and there exists a continuously differentiable function $V : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\alpha_1|\zeta|^2 \leq V(\zeta,t) \leq \alpha_2|\zeta|^2\,, \tag{F.4}$$

and for each $z \in C^0$,

$$\frac{\partial V}{\partial \zeta}A(z,t)\zeta + \frac{\partial V}{\partial t} \leq -\alpha_3|\zeta|^2 - \alpha_4\left(\frac{\partial V}{\partial \zeta}E(z,t)\right)^2 \tag{F.5}$$